

## University of Groningen

### Over de zin van opbrengstgericht(er) werken in het onderwijs

Visscher, A.J.

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2015

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Visscher, A. J. (2015). *Over de zin van opbrengstgericht(er) werken in het onderwijs*. GION onderzoek/onderwijs.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



rijksuniversiteit  
 groningen

faculteit gedrags- en  
 maatschappijwetenschappen

Adrie Visscher

oratie

## **Over de zin van opbrengstgericht(er) werken in het onderwijs**

**Prof. dr. A.J. Visscher**

# **Over de zin van opbrengstgericht(er) werken in het onderwijs**

## **Rede**

op 8 december 2015

in verkorte vorm uitgesproken

bij de aanvaarding van het ambt van bijzonder hoogleraar

*Opbrengstgericht werken in het onderwijs*

aan de Faculteit der Gedrags- en Maatschappijwetenschappen

van de Rijksuniversiteit Groningen



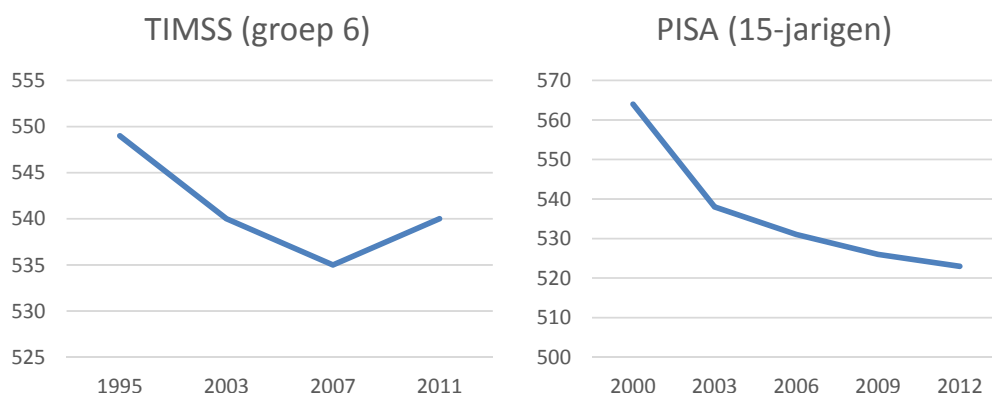
Door

**Prof. dr. A.J. Visscher**

Leden van het College van Bestuur, zeer geachte aanwezigen,

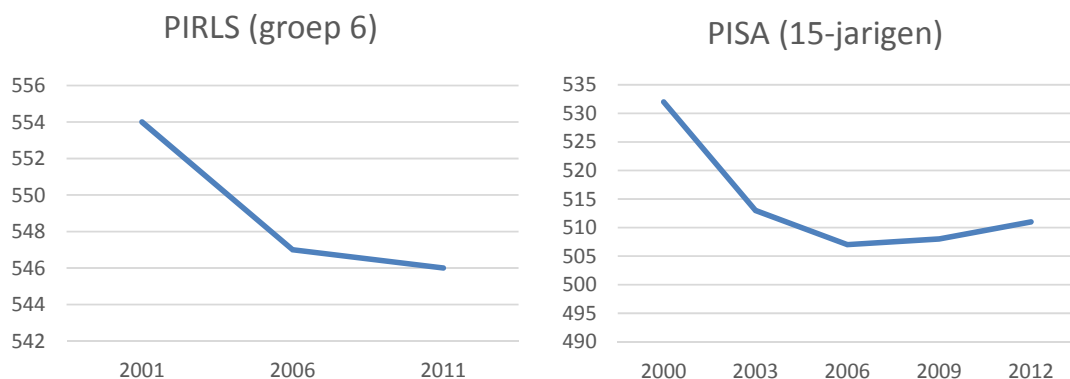
‘Opbrengstgericht werken’ (afgekort als OGW) is een specifiek Nederlandse term die in 2007 (Ministerie voor OC&W, 2007) door de Nederlandse overheid werd geïntroduceerd naar aanleiding van de veronderstelde achteruitgang van de prestaties van Nederlandse leerlingen in de kernvakken in internationale vergelijkingen als PISA, PIRLS en TIMSS.

De overheid maakte zich zorgen en wilde verbetering van de prestaties van Nederlandse leerlingen (Visscher en Ehren, 2011). In het onderwijs zou daarom vooral in de kernvakken (rekenen, taal, lezen) opbrengstgericht gewerkt moeten worden. Dat wil zeggen dat data over met name *de prestaties* van leerlingen en vooral verkregen uit *leerlingvolgsystemen* meer benut zouden moeten worden. Dit zou moeten leiden tot een meer *doelgerichte* werkwijze, én tot het bieden van meer *onderwijs op maat*, oftewel tot onderwijs dat past bij de, met behulp van genoemde data, vastgestelde *beheersingsniveaus* en *onderwijsbehoeften* van leerlingen. Op die manier zou systematisch gewerkt kunnen worden aan de maximalisatie van de prestaties van zowel hoog, gemiddeld, als relatief laag presterende leerlingen (Inspectie van het onderwijs, 2010, 2015; Ministerie voor OCW, 2010).



Figuur 1: TIMSS- en PISA-prestaties rekenen/wiskunde

In de twee bovenstaande grafieken (Figuur 1) ziet u de gemiddelde scores van Nederlandse leerlingen voor het vak *rekenen/wiskunde*, links voor leerlingen uit groep 6 van de basisschool, rechts voor 15-jarigen in het voortgezet onderwijs. Het gaat hier om twee verschillende studies, met twee onvergelykbare schalen (dit geldt ook voor Figuur 2). De TIMSS-grafiek laat zien dat de rekenprestaties van leerlingen in groep 6 in de TIMSS-studie in de periode 1995 - 2007 achteruit gingen om vervolgens weer iets te herstellen. De PISA-grafiek toont de daling van de leerlingprestaties van 15-jarigen in de periode 2000 - 2006 (en daarna).



Figuur 2: PIRLS- en PISA-prestaties begrijpend lezen

Bij *begrijpend lezen* (Figuur 2) zien we eveneens dalingen in de PIRLS- en PISA-studies tijdens de periode 2000 - 2006.

Cijfers over de onderwijsprestaties van landen in vergelijking met die van andere landen zijn in meerdere opzichten *relatief*. Als het onderwijs in de landen waarmee we ons vergelijken sterk voor verbetering vatbaar is en men scoort als land hoog te midden van die groep landen, dan houdt men zichzelf voor de gek wanneer men concludeert dat men op niveau presteert. Het omgekeerde geldt natuurlijk ook: niet in de top presteren te midden van landen waarin het onderwijs heel goed is betekent niet dat men een slecht onderwijssysteem heeft, maar dat het *nóg* beter kan. Hoewel het hier te ver voert om daar op in te gaan is een belangrijke vraag dus van welke kwaliteit het onderwijs is in de landen waarmee we ons vergelijken.

Los daarvan is maar liefst 1,3 miljoen Nederlanders tussen de 16 en 65 jaar laaggeletterd, oftewel meer dan 10% van alle Nederlandse 16 - 65 jarigen! Hoewel laaggeletterden wel een beetje kunnen lezen en schrijven hebben ze daar grote moeite mee hetgeen een enorme impact heeft op hun maatschappelijk functioneren (43% van hen is bijvoorbeeld werkloos) én op hun persoonlijk welbevinden. Het gaat hier, tegen de verwachting in, niet voornamelijk om allochtonen; 65% van de laaggeletterden is autochtoon. Dit *absolute* aantal laaggeletterden is naar mijn mening, los van de omvang van dit probleem in andere landen, op zich voldoende reden om in Nederland naar verbetering van de onderwijsprestaties te streven.

Zoals ik al aangaf verwachtte de overheid veel van het gebruik van leerlingvolgsystemen. Om een indruk te geven van de informatie die leerlingvolgsystemen kunnen bieden laat ik hier drie voorbeelden zien van analyses die men met het Cito-leerlingvolgsysteem kan uitvoeren. De eerste analyse betreft de categorieënanalyse waarmee men inzicht krijgt in de mate waarin elke leerling de diverse *onderdelen* van de leerstof van een vak beheerst. In Figuur 3 ziet u hoe de leerlingen bij rekenen/wiskunde scoren op getalbegrip (GET), optellen en aftrekken (O&A), vermenigvuldigen en delen (V&D), en meten/tijd/geld (MTG). Deze informatie kan de leerkracht gebruiken bij het nemen van instructiebeslissingen. Elke leerling heeft een toetsscore en een bijbehorend prestatieniveaulabel: A (de 25 procent hoogst presterende leerlingen), B, C, D, dan wel E (de 10 procent laagst presterende leerlingen). Aangegeven wordt ook of een leerling in positieve dan wel in negatieve zin op de onderdelen van een vakgebied opvalt gegeven zijn/haar overall toetsscore. Zo scoort de fictieve leerlinge Carlijn Wibbershoven voor een leerling met een toetsscore van 55 opvallend laag op

Getalbegrip (-18) en op Meten, Tijd en Geld (-38) terwijl ze juist opvallend hoog scoort op Optellen en Aftrekken (+7) en ook op Vermenigvuldigen en Delen (+24). Voor Bart Vissers geldt dat zijn scores *niet opvallend* afwijken van het patroon dat op basis van zijn overall score van 56 werd verwacht.

Groep: 5-5B

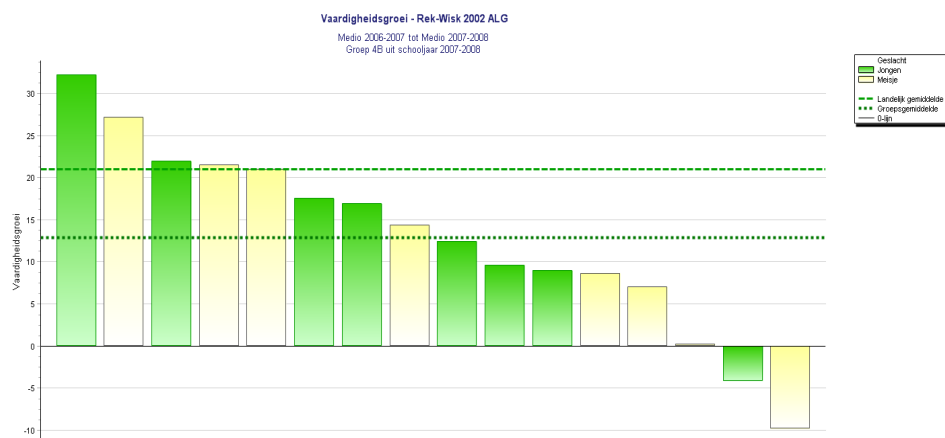
Toets-taak: Rek-Wisk 2009 -E4

	Toets score		Signaal	GET		O&A		V&D		MTG	
				s%	afw	s%	afw	s%	afw	s%	afw
Veerle Lucassen	56	C	Niet opvallend	78	+3	64	-2	71	+10	48	-21
Lejla van Motmar	43	D	Opvallend	65	+8	38	-4	48	+15	12	-32
Ivo Notemans	64	B	Niet opvallend	84	0	85	+7	76	-1	73	-8
Sheswin Ozturk	52	C	Niet opvallend	73	+2	51	-8	59	+6	61	-1
Bart Vissers	56	C	Niet opvallend	78	+2	57	-9	73	+11	58	-12
Carlijn Wibbershoven	55	C	Zeer opvallend	57	-18	72	+7	84	+24	30	-38
Jordy Wiekken	63	B	Niet opvallend	86	+3	74	-3	73	-3	88	+7

s% = Percentage score geobserveerd. Afw = Percentage score afwijking

Figuur 3: De categorieënanalyse

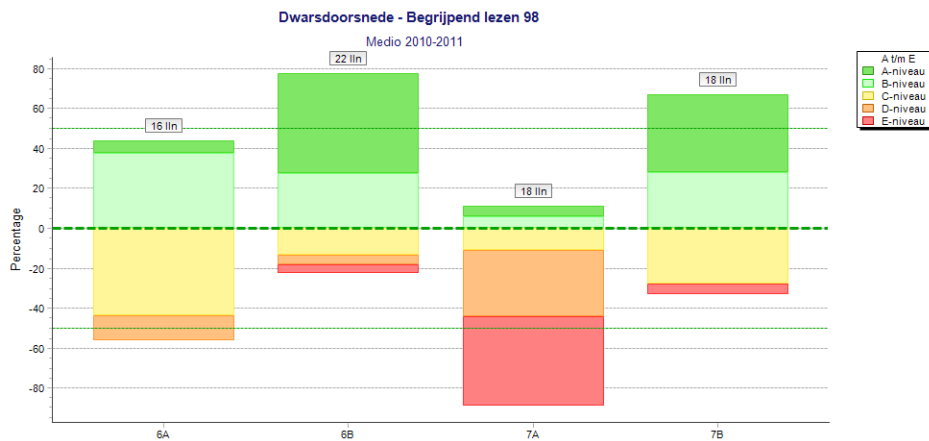
Onderstaande Figuur 4 geeft per leerling (elke staaf representeert een leerling) de vaardigheidsgroei tussen twee toetsen weer voor het vak rekenen/wiskunde evenals hoe deze groei zich verhoudt tot het gemiddelde van de groep (de onderste stippellijn) en tot het landelijk gemiddelde (de bovenste stippellijn). Het is informatief om ook naar de vaardigheidsgroei te kijken in plaats van alleen naar een *vaardigheidsscore op één moment*. Een leerling met een lage vaardigheidsscore kan sterk gegroeid zijn sinds de vorige toets terwijl een leerling met een hoge vaardigheidsscore helemaal niet vooruit kan zijn gegaan (de twee leerlingen uiterst rechts zijn niet gegroeid, maar zelfs in vaardigheid 'gekrompen' in vergelijking met de vorige toets). De groei van elke leerling kan vergeleken worden met de landelijk gemiddelde groei van *vergelijkbare* leerlingen (bijvoorbeeld met de 25% hoogst presterende, of de 10% laagst presterende leerlingen). Die groei verschilt namelijk tussen A-, B-, C-, D-, E-leerlingen.



Figuur 4: De vaardigheidsgroei per leerling tussen twee leerlingvolgysteemtoetsen

De dwarsdoorsnede in Figuur 5 is relevanter voor de evaluatie van de resultaten op *schoolniveau* (en eventueel ook op *bestuursniveau*). Elke rechthoek vertegenwoordigt een groep (de aantallen leerlingen staan per groep aangegeven). Per groep geven de kleuren weer welk deel van de leerlingen in een groep op een bepaald niveau (A tot en met E) presteert. Een school kan hier snel zien hoe elk

van haar groepen presteert in vergelijking met de landelijke benchmark van 50% in totaal op A- en B-niveau presterende leerlingen, en 50% in totaal op C-, D- en E-niveau presterende leerlingen.



Figuur 5: De dwarsdoorsnede

Nu de aanleiding van het OGW-beleid geschetst is zal ik eerst ingaan op de inhoud van het begrip 'Opbrengstgericht Werken'. Vervolgens zal ik de evidentie met betrekking tot de effectiviteit van OGW bespreken. Daarna zal ik aangeven hoe toekomstig OGW-onderzoek op dit terrein er naar mijn mening idealiter uitziet, om vervolgens af te sluiten met enkele aanbevelingen ter bevordering van OGW in de onderwijspraktijk.

## 1. Het begrip Opbrengstgericht werken

Wat de term OGW betreft geven Blok en collega's (2015) in een recent artikel in *Pedagogische Studiën* terecht aan dat opbrengstgerichtheid inherent is aan het onderwijs. Hoewel elke sector werknemers kent die meer weekendgericht dan opbrengstgericht werken, werkt men, wanneer men werkt, doorgaans aan de realisatie van de opbrengsten waarvoor de organisatie is opgericht. Zo wordt in Nederlandse scholen ook hard gewerkt aan leeropbrengsten. In die zin is opbrengstgericht werken taalkundig gezien net zo'n vreemde term als 'bestemmingsgericht reizen', 'oplossingsgericht puzzelen' en 'competentieverricht leren'. Om welke andere reden, dan het verwerven van bepaalde kennis, een bepaalde attitude en/of een vaardigheid zou iemand immers leren? Blok et al. (2015) stellen daarom 'opbrengstgerichtER werken' voor als meer passende term. Daarmee ben ik het eens, ook omdat het heel nuttig is om de opbrengstgerichtheid binnen het onderwijs te *versterken*, in de zin van meer expliciet resultaatgericht werken aan de maximalisatie van onderwijsopbrengsten.

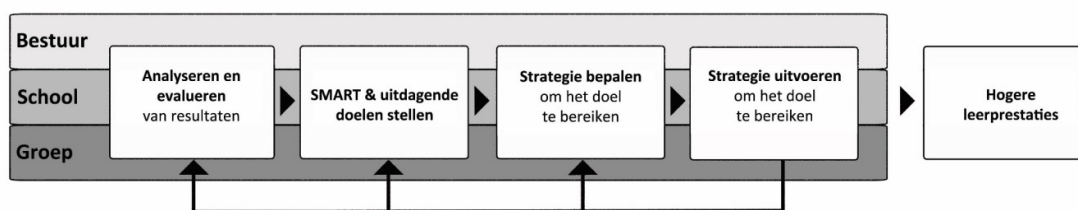
De kern van opbrengstgericht werken op het niveau *van de klas* wordt naar mijn mening goed getypeerd door Dylan Wiliam (die het dan overigens heeft over 'assessment for learning', dat binnen OGW een belangrijke rol speelt): "Using evidence of achievement to adapt what happens in classrooms to meet learner needs." (2006). De evaluatie van welke leerlingen welke leerstofonderdelen wel/niet beheersen, vormt hier de basis voor het zoveel mogelijk afstemmen van de instructie in brede zin (uitleg leerstof, het geven van opdrachten, het gebruik van leermaterialen en het ter beschikking stellen van een bepaalde hoeveelheid tijd om aan iets te werken) op wat elke leerling nodig heeft. Een opbrengstgerichte werkwijze gaat echter verder dan alleen leerlinggerichte maatregelen *in de klas*.



Maatregelen op *school- en bestuursniveau* kunnen ook belangrijk zijn in zo'n werkwijze. Het gaat dan op die niveaus om maatregelen op basis van data, die bedoeld zijn om *indirect* onderwijskwaliteit en leerprestaties te maximaliseren. Een school zou bijvoorbeeld n.a.v. tegenvallende prestaties kunnen signaleren dat de effectieve leertijd in de school beperkt is, en daarom een maatregel kunnen nemen om die leertijd schoolbreed te vergroten. Een bestuur zou een schoolleider die jaarlijks in positieve zin opvalt door zijn datagebruik en hoge schoolprestaties kunnen vragen om zijn werkwijze te delen met zijn collega-schoolleiders, opdat de collega's een dergelijke werkwijze ook meer gaan volgen.

Onderstaande Figuur 6 geeft weer dat er bij OGW op klas-, school- en bestuursniveau steeds sprake is van dezelfde, rationele aanpak en stappen. Steeds vormen de evaluatie, de analyse en de interpretatie van leerlingprestatiegegevens het uitgangspunt voor het stellen van doelen die SMART<sup>1</sup> en uitdagend zijn. Wanneer de huidige situatie (het eerste blok in Figuur 6) en de gewenste situatie (het tweede blok) bekend zijn, dan kan een strategie bepaald worden die men kansrijk acht om het gestelde doel (het derde blok) te realiseren. Het vierde blok werd in de beleidsnota's van de overheid niet genoemd. Het nadenken over en bepalen van een goede probleemaanpak werd impliciet gelijkgesteld aan de uitvoering daarvan. De praktijk leert echter dat dit niet realistisch is. In het gedicht *Het Huwelijk* heeft Elschot het over de echtgenoot die het fantastisch lijkt om na de vele huwelijksjaren zijn vrouw dood te slaan en het huis in brand te steken. *“Maar doodslaan deed hij niet, want tussen droom en daad staan wetten in de weg en praktische bezwaren.”* Een goed plan is met andere woorden mooi, maar een uitgevoerd goed plan is nog veel mooier en voor veel plannen geldt dat ze de uitvoeringsfase nooit bereiken. In blok vier in Figuur 6 wordt een plan op één van de drie niveaus waarop men opbrengstgericht kan werken in praktijk gebracht, en leidt dit idealiter tot betere leerprestaties.

Hoewel de overheid in haar beleidsstukken het gebruik bepleit van prestatiedata uit leerlingvolgsystemen, beperkt men zich daartoe bij OGW idealiter niet, maar benut men alle relevante informatie over leerlingen, klassen en scholen waarover men kan beschikken op basis van observaties in de klas, andere afgenomen toetsen (methodetoetsen bijvoorbeeld), diagnostisch onderzoek en op andere manieren wijze verkregen (achtergrond)informatie.



Figuur 6: Componenten en niveaus van OGW (gebaseerd op Keuning et al., in druk)

<sup>1</sup> Er bestaan uiteenlopende definities voor sommige letters van dit acroniem, ik ga hier uit van de volgende definitie: specific, measurable, achievable, relevant, time bound.

Gegeven het voorgaande stel ik de volgende definitie van OGW voor:

Opbrengstgericht werken betekent dat leerkrachten, scholen en besturen een systematische werkwijze hanteren waarbij:

- regelmatig, op basis van toetsgegevens en andere relevante gegevens, de vorderingen van leerlingen, groepen en scholen evenals de oorzaken en achtergronden van die vorderingen worden bepaald, en waarbij
- op basis van de verkregen inzichten op klas-, school- en bestuursniveau doelen worden geformuleerd en daarbij passende maatregelen worden bepaald en uitgevoerd met het oog op het maximaal presteren van alle leerlingen.

Deze definitie past goed bij de internationaal gangbare term ‘data-based decision making’ (Ikemoto en Marsh, 2007) waarbij het ook gaat om het in brede zin benutten van data en het in brede zin nemen van maatregelen om onderwijskwaliteit en leerprestaties te optimaliseren: “teachers, principals, and administrators systematically collecting and analyzing data to guide a range of decisions to help improve the success of students and schools” (p108). Wel dient hierbij opgemerkt te worden dat Ikemoto en Marsh alleen over ‘decisions’ spreken en niet over de implementatie van die beslissingen.

## **2. De effectiviteit van Opbrengstgericht Werken**

Ik ga hier eerst in op wat bekend is over de condities waaronder de twee kernprincipes van OGW, feedback<sup>2</sup> en goal setting (Visscher & Ehren, 2011), effectief zijn. Na deze meer theoretische uiteenzetting zal ik een overzicht geven van de resultaten van empirisch onderzoek naar de effecten op leerprestaties van zes Nederlandse OGW-interventies waarin feedback en goal setting op een bepaalde manier zijn ingevuld.

### **2.1 Factoren die de effecten van OGW beïnvloeden**

Op basis van meer dan 100 jaar feedback-onderzoek evenals uitgebreid onderzoek naar goal setting is er sterk wetenschappelijk sterk bewijs dat feedback (Kluger en DeNisi, 1996; Hattie en Timperley, 2007; Visscher en Coe, 2002) en goal setting (Locke en Latham, 2002) sterk prestatieverbeterend *kunnen* werken. Aangetoond is dat wanneer feedback en goal setting gecombineerd worden, de effecten nog sterker zijn, dan de som van de effecten van elk van hen (Locke en Latham, 2002). De verklaring hiervoor is dat men wanneer men regelmatig feedback krijgt over de mate waarin gestelde doelen bereikt worden, men, wanneer dit wenselijk is, minder productieve werkwijzen kan vervangen door (hopelijk) effectievere.

Een concreet voorbeeld van een schoolverbeteringsprogramma waarin de principes van feedback en goal setting op effectieve wijze zijn ingebouwd is het programma *Success for All* waarin het frequent meten van leerlingprestaties de basis vormt voor het regelmatig (bij)stellen van prestatiedoelen en het

---

<sup>2</sup> In veel studies en publicaties staat feedback aan *leerlingen* centraal, in deze oratie gaat het vooral om feedback aan *leerkrachten* (en ook aan scholen en besturen).

samenstellen van homogene leesgroepen waarin leerlingen dagelijks negentig minuten op hun niveau leesles krijgen.

Hoewel feedback en goal setting sterk prestatieverbeterend *kunnen* werken weten we ook dat dit niet altijd en overal het geval is. Kluger en DeNisi (1996) vonden in hun beroemde meta-analyse een gemiddelde effectgrootte van 0,41, maar ook dat de 607 effectgroottes in de 131 onderzochte studies sterk varieerden. Ondanks het overall positieve effect van 0,41 bleek feedback in 38% van de gevallen zelfs een *negatief* effect op prestaties te hebben (het ging hier niet alleen om feedback in onderwijscontexten)! Het hangt dus van een aantal condities af of feedback wel of niet tot de gewenste effecten leidt. Hetzelfde geldt voor goal setting. Zo blijkt de aard van de taken waarover feedback wordt verkregen en waarvoor doelen worden gesteld een belangrijke factor te zijn. Bij complexe taken worden minder grote prestatie-effecten gevonden, dan bij eenvoudige. De oorzaak daarvoor wordt waarschijnlijk gevormd door het feit dat het bij complexe taken moeilijker is om prestatieverbeteringsstrategieën te bepalen en uit te voeren indien de feedback laat zien dat gestelde doelen niet behaald worden. Onderwijs geven is allesbehalve simpel en als feedback laat zien dat je als leerkracht of school onder het landelijk gemiddelde scoort met een bepaald vak, dan is daarmee de 'remedie' nog allerm minst gegeven.

Er zijn meer factoren die bepalen of feedback en goal setting werken. Ik zal eerst ingaan op factoren die belangrijk zijn voor de effecten van goal setting en vervolgens op condities voor effectief feedbackgebruik.

#### *2.1.1 Goal setting, hoe werkt het?*

Het stellen van SMART en uitdagende doelen blijkt gemiddeld tot betere prestaties te leiden, doordat dit leidt tot doelgericht gedrag (en dus tot de reductie van die activiteiten die irrelevant zijn voor doelbereiking), een grote inzet om de gestelde doelen te realiseren evenals het zoeken naar effectieve strategieën voor doelrealisatie (Locke & Latham, 2002). Echter, goal setting is geen wondermiddel. Zo zijn de effecten van het stellen van doelen afhankelijk van de mate waarin men gemotiveerd is voor de gestelde doelen (goal commitment), en, zoals reeds werd aangegeven, de mate waarin de te verrichten taken complex zijn. Bij relatief simpele taken zal het eenvoudiger zijn om met extra inzet en focus betere prestaties te realiseren en bij complexe taken is het vinden van effectieve strategieën voor doelrealisatie niet minder essentieel, maar moeilijker. Tot slot zijn de competenties van degene die de doelen stelt van invloed op de mate waarin het stellen van doelen tot betere prestaties leidt. In het geval van OGW zijn de competenties van leerkrachten, schoolleiders en ook bestuurders dus belangrijk voor het werken met goal setting.

#### *2.1.2 Feedbackkenmerken die ertoe doen*

Het voert hier te ver om een uitputtend overzicht te presenteren van alle feedbackkenmerken die een rol spelen bij de benutting van feedback. Ik bespreek daarom die factoren die het belangrijkste zijn voor ons thema.

Feedback moet allereerst *toegankelijk* zijn voor de gebruiker. Leerlingvolgysteemoutput is niet super toegankelijk, de correcte interpretatie ervan vereist scholing. Met die scholing blijkt het correcte gebruik van deze vorm van feedback echter goed te leren te zijn (Staman et al., 2014).

Verder is de *inhoud van de feedback* belangrijk voor de effecten die het gebruik van feedback zal sorteren. Zo kan feedback meer of minder *handvatten voor verbetering* bieden. Zoals gezegd vonden Kluger en DeNisi in hun meta-analyse dat een derde van de geanalyseerde effecten negatief was. De feedback bleek de ontvanger in deze gevallen geen indicatie te geven van hoe deze zijn/haar presteren kon verbeteren. Vertaald naar het onderwijs kan hierbij bijvoorbeeld gedacht worden aan het door de leerkracht louter ontvangen van een prestatiescore per leerling als feedback. Een leerkracht heeft voor het nemen van instructiebeslissingen veel meer aan informatie over welk leerstofonderdeel (bijv. getalbegrip, optellen/afrekken, etc.) door bepaalde leerlingen wel/niet beheerst wordt, in combinatie met diagnostische informatie (welk type fout maakt men en waardoor?) en idealiter ook in combinatie met tips over hoe in de instructie op het prestatiebeeld van leerlingen zou kunnen worden aangesloten.

Andere belangrijke feedbackaspecten betreffen de *frequentie* en de *tijdigheid* van de feedback (Shute, 2008; Faber & Visscher, 2014). Zo kan men zich afvragen of de feedback uit een leerlingvolgsysteem die slechts twee keer per jaar wordt verkregen (waarvan één keer als het schooljaar ten einde is en de leerlingen vaak naar een andere leerkracht gaan) niet erg weinig is en niet (te) laat komt. Populaire digitale oefensystemen, zoals Snappet, Rekentuin en Taalzee bieden direct *tijdens de les* feedback aan de leerkracht<sup>3</sup> (welke opgaven blijken voor de leerlingen problemen op te leveren?) waardoor de leerkracht een probleem snel na het ontstaan kan aanpakken.

Feedbackbenutting vereist ook dat de leerkracht de feedback kan benutten voor het bieden van instructie op maat. Hiermee zijn we aangekomen bij een tweede belangrijke factor voor de effecten van feedback in het kader van OGW.

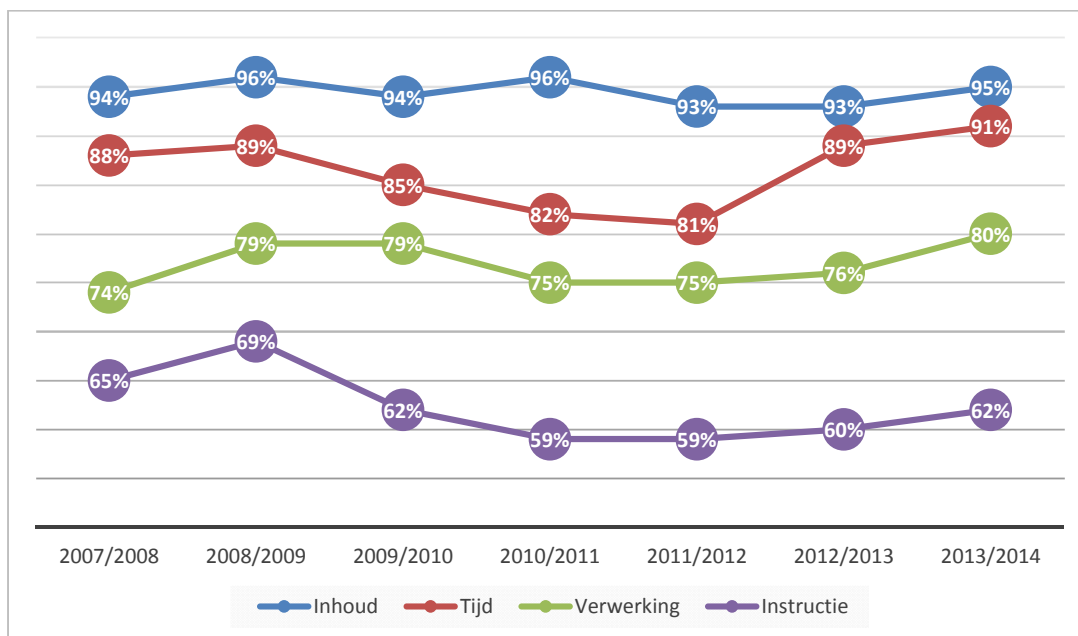
### 2.1.3 De feedbackontvanger als beïnvloedende factor

Belangrijk voor de effecten van OGW is hoe met de feedback wordt omgegaan. Dit zal vooral afhankelijk zijn van de *competenties van de ontvanger*: in hoeverre vindt de ontvanger de prestatiefeedback waardevol en kan hij/zij die correct interpreteren en op basis daarvan adequate maatregelen bepalen en uitvoeren? In het geval van een leerkracht zal dit onder andere afhangen van zijn/haar evaluatieve competenties (bijvoorbeeld leerlingvolgsysteemoutput correct kunnen interpreteren, kennis van de kenmerken van verschillende toetsen), de pedagogische en vakinhoudelijke kennis (pedagogical content knowledge) en zijn/haar algemene didactische competenties (o.a. lesopbouw, duidelijke instructie, klasmanagement). Vooral de vaardigheid om de instructie op basis van de ontvangen prestatiefeedback adequaat aan te kunnen passen en te kunnen differentiëren is van belang, maar daarvoor dienen ook andere didactische competenties (bijvoorbeeld klasmanagement) beheerst te worden, want bijvoorbeeld in een ongeorganiseerde klas komt men niet aan differentiëren toe.

In ieder geval een deel van de voor OGW benodigde kennis en vaardigheden zal zonder specifieke scholing niet voldoende aanwezig zijn binnen scholen, bijvoorbeeld de kennis en vaardigheden die nodig zijn voor de correcte analyse van leerlingvolgsysteemgegevens evenals differentiatiecompetenties (Inspectie van het Onderwijs, 2014). Onderstaande Figuur 7 laat zien in hoeverre leerkrachten in het basisonderwijs volgens de Onderwijsinspectie de leerinhoud, tijd, verwerking en instructie differentiëren (Inspectie van het Onderwijs, 2015). Vooral bij het differentiëren van de instructie valt veel winst te behalen.

---

<sup>3</sup> En ook aan leerlingen.



Figuur 7: Leerkrachtscores voor vier vormen van differentiatie (Inspectie van het Onderwijs, 2015)

Pabo's slagen er nog onvoldoende in om leerkrachten met deze competenties toe te rusten (terwijl het leerlingvolgsysteem toch al ruim twintig jaar bestaat en vrijwel elke basisschool ermee werkt). Daar bestaat dus ook veel ruimte voor verbetering.

Ook op school- en bestuursniveau zal opbrengstgericht werken een beroep doen op nieuwe competenties, bijvoorbeeld de competenties voor het jaarlijks stellen van heldere, uitdagende maar ook haalbare schoolprestatiedoelen (dit is moeilijker, dan het op het eerste oog lijkt), het als schoolteam monitoren van de mate van doelrealisatie, het evalueren van de werkwijzen die wel/niet gewerkt hebben en het op basis daarvan leren en verbeteren. Het is begrijpelijk dat schoolleiders met de introductie van OGW niet simpelweg ineens volledig op een resultaatgerichte manier (kunnen) gaan werken met hun teamleden. Dit zal in veel scholen een culturomslag en een geleidelijk proces van schoolontwikkeling vereisen. Met scholing en begeleiding is hier ook veel te winnen.

#### 2.1.4 De invloed van de schoolcontext

De *schoolcontext* waarbinnen OGW wordt ingevoerd speelt eveneens een belangrijke rol. We weten dat die context varieert tussen scholen (Maslowski et al., 2015; Deunk et al., 2015; Van Geel, Teunis & Visscher, 2015). Eenpitter-besturen werken doorgaans met vrijwilligers als bestuurders die niet altijd over onderwijskundige deskundigheid beschikken. Kleine scholen hebben vaak parttime schoolleiders en intern begeleiders, die minder tijd aan OGW kunnen besteden terwijl een omvattende innovatie à la OGW zowel in grote als in kleine scholen een basisinvestering van een bepaalde omvang vereist (Maslowski et al., 2015). Ook verschillen scholen wat betreft de kwaliteit van hun leerkrachten en intern begeleiders, de mate van onderwijskundig leiderschap en tot slot ook qua personeelsmobiliteit. In de uitgevoerde OGW-projecten bleek deze mobiliteit van zowel leerkrachten als schoolleiders niet gering en slecht te zijn voor de implementatie van OGW (Van Geel, Teunis & Visscher, 2015). De reden

daarvoor is waarschijnlijk dat nieuwe schoolleiders en leerkrachten niet betrokken waren geweest bij de beslissing om aan OGW te gaan werken. Een nieuwe schoolleider zal niet per se het OGW-initiatief steunen. Ook ontberen de nieuwkomers waarschijnlijk bepaalde OGW-kennis die hun collega's zich al eigen hebben gemaakt tijdens de OGW-interventie. Bovendien hebben nieuwe werknemers tijd nodig om zich in te werken in hun nieuwe werksituatie waardoor men minder tijd zal hebben om aan OGW te werken.

Het samenspel van genoemde schoolcontextfactoren leidt tot tussen scholen variërende implementatiecontexten, die zich meer of minder lenen voor de invoering van OGW. Eén van de uitdagingen bij het bevorderen van OGW is het omgaan met deze variatie tussen scholen. Een belangrijke vraag daarbij is of een 'one size fits all' implementatiestrategie, die naar tempo gedifferentieerd wordt aangeboden voldoende is, of dat bepaalde contexten specifiek op die contexten toegesneden interventie-aanpakken vereisen. Daarmee zijn we aangekomen bij de vierde en laatste factor die van belang is voor de benutting van feedback (en voor de effectiviteit van OGW), namelijk de aard van de professionalisering voor OGW.

#### *2.1.5 Professionalisering als noodzakelijke voorwaarde voor OGW*

Gegeven de complexiteit van OGW zal de aard van de scholing en de begeleiding van leerkrachten, schoolteams en hun bestuurders sterk bepalend zijn voor of OGW 'gaat werken' of niet (Visscher & Ehren, 2011). Professionaliseringstrajecten moeten voorzien in de nog ontbrekende competenties die OGW veronderstelt. Black en Wiliam (1998) komen op basis van hun bekende meta-analyse betreffende de benutting van formatieve toetsgegevens door leerkrachten tot de conclusie dat dit type veranderings- en professionaliseringsprocessen heel geleidelijk verlopen en daarom een kwestie van lange adem zijn. Slavin et al. (2011) vonden dat leerkrachten pas in het derde jaar van hun data-based decision-making interventie hun lessen aan pasten op basis van leerlingprestatiedata. De eerste twee jaren beperkte men zich vooral tot het analyseren van leerlingprestatiedata: "...first and second year interventions were analogous to taking a patient's temperature, not providing a treatment" (Slavin et al., 2013, p. 22).

Een belangrijke vraag is waaraan een professionaliseringstraject moet voldoen om OGW succesvol in te kunnen voeren. Wat leert de literatuur ons hierover? De literatuur over de in-service professionalisering van leerkrachten biedt slechts beperkt aangrijpingspunten voor de inrichting van OGW-professionaliseringstrajecten. Er is nog opmerkelijk weinig onderzoek beschikbaar waarin de effecten van varianten van leerkrachtprofessionalisering (in algemene zin, niet specifiek voor OGW) experimenteel onderzocht zijn. Yoon et al. (2007) vonden voor de periode 1986 -2003 wereldwijd slechts 9 studies met een experimenteel design. Deze experimenten wezen vooral op het belang van een behoorlijke trainingsomvang (vanaf plus minus 50 uur werden positieve effecten gevonden, hoewel dat natuurlijk niet zoveel zegt, los van bijvoorbeeld de doelstelling van de professionaliseringstrajecten). Het niet-experimenteel onderzoek biedt vooral nader te verifiëren *indicaties* dat bepaalde factoren belangrijk zijn bij professionalisering.

De literatuur over in-service leerkrachtprofessionalisering is nogal globaal. Er wordt in deze literatuur onder andere gewezen op het belang van het werken op basis van een heldere verbeteringstheorie (Desimone, 2009; zie Figuur 8). Voor OGW betekent dit dat een goed doordacht plan nodig is met

betrekking tot hoe een OGW-training in gespecificeerde, voor OGW nodig geachte competenties zal resulteren en hoe dát weer zal leiden tot ander onderwijs in de klas, en daardoor tot betere leerlingprestaties.

Ook wordt gewezen op het belang van factoren in de schoolorganisatie en schoolcontext, bijvoorbeeld onderwijskundig leiderschap, een veilig innovatieklimaat, motivatie voor de vernieuwing, collegiale samenwerking (Vescio et al., 2008) evenals de schoolinterne facilitering van innovatie-activiteiten (Timperley 2008; van Veen et al., 2011). In deze literatuur wordt ook gewezen op een aantal globale leerpsychologische principes waarmee rekening dient te worden gehouden. Zo wordt het belang van ‘*deep learning*’ benadrukt: de integratie van kennis en vaardigheden, het als leerkracht begrijpen waaróm een bepaalde instructie-aanpak belangrijk is, i.p.v. het mechanistisch uitvoeren van zo’n aanpak (Timperley, 2008). Ook wordt professionalisering aan de hand van een *coherente set gevarieerde leeractiviteiten* bepleit, bijvoorbeeld de combinatie van face-to-face leren, leren toepassen, feedback krijgen en reflecteren op het eigen functioneren.

De literatuur over de scholing van in-service leerkrachten is óók globaal in die zin, dat er geen onderscheid wordt gemaakt in *de aard van* de te leren skills, competenties en attitudes. Het lijkt alsof het niet uitmaakt wat geleerd moet worden, altijd zouden dezelfde professionaliseringsfactoren belangrijk zijn. Sommige kennis kan echter goed door middel van een paar uur mondelinge uitleg worden overgebracht terwijl bijvoorbeeld het leren beheersen van complexe didactische vaardigheden intensieve begeleidingstrajecten (ook in de klas) vereist.



Figuur 8: Theory of improvement voor leerkrachtprofessionalisering (gebaseerd op Desimone, 2009)

Hoewel de professionaliseringsliteratuur over de factoren waarmee men in professionaliseringstrajecten rekening moet houden relevant is, blijft ongewis hoe men een (OGW-) professionaliseringsinterventie nu precies het beste vorm kan geven. Hoe kan men bijvoorbeeld zorgen voor een ‘gevarieerd en coherent geheel aan leeractiviteiten’ (en wanneer is er bijvoorbeeld wel/geen sprake van coherentie?), voor sterk onderwijskundig leiderschap en voor de beheersing van een complexe vaardigheid als differentiëren?

Tijdens professionaliseringstrajecten wordt in het onderwijs veel gewerkt op basis van mondelinge kennisoverdracht, bijvoorbeeld het mondeling overbrengen van informatie over hoe nieuw curriculummateriaal gebruikt kan worden. Het daadwerkelijk leren toepassen van nieuwe werkwijzen en vaardigheden in de klas krijgt veel minder aandacht. “*The history of instructional interventions has a thin record of success*”, aldus Cohen & Ball (2001). Dat komt volgens de auteurs onder andere, doordat veel initiatieven om de instructie te verbeteren een *partieel* karakter kennen: zelden is er sprake van een geïntegreerde benadering van hetgeen een bepaalde vernieuwing vereist van scholen, leerkrachten, leerlingen, leermaterialen en de organisatorische context waarin dit alles moet plaatsvinden. Veel wordt aan leerkrachten overgelaten, zij moeten zelf veel van de, veelal globaal omschreven, vernieuwing invullen, bijvoorbeeld hoe met nieuwe lesmaterialen gewerkt moet worden.

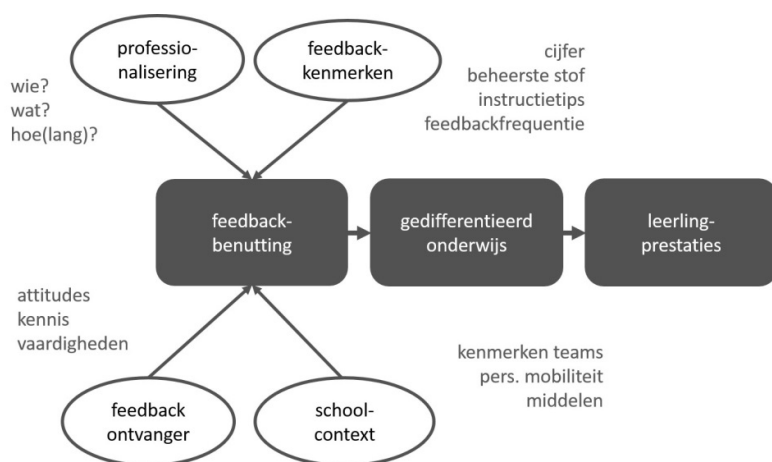
Dat leidt tot tal van eigen, uiteenlopende interpretaties, die meer of minder zullen stroken met het beoogde gedrag. Omdat het beoogde nieuwe leerkrachtgedrag niet helder omschreven is, worden leerkrachten daar ook niet systematisch in getraind en begeleid (Slavin, 2015). De kans is daardoor groot dat er in de klas niets wezenlijks verandert en dat daardoor de gewenste effecten op leerprestaties niet gerealiseerd worden.

De literatuur over het ontwerpen van instructie voor *het verwerven van complexe professionele vaardigheden* (inclusief kennis en attitudes) voor uiteenlopende situaties waarin sprake is van complex leren is aanzienlijk specifiek, dan de algemene wetenschappelijke literatuur over de professionalisering van leerkrachten. Specifieker in de zin van een uitgewerkte *methodiek* om de inhoud van trainingen voor complex leren te bepalen en eveneens specifiek wat betreft de leerpsychologische uitgangspunten voor het leren van complexe vaardigheden. In het 4-componenten instructional design model van Van Merriënboer & Kirschner (2013) wordt gestart met een cognitieve taakanalyse. Voor OGW op leerkrachtniveau zou dat bijvoorbeeld betekenen dat gestart wordt met een analyse van de te verrichten OGW-taken, en de kennis, vaardigheden en attitudes die daarvoor essentieel zijn. Uit welke routines en niet-routinematige activiteiten bestaan die OGW-taken, zoals de analyse van data en het in de klas bieden van instructie op maat op basis van geanalyseerde data? En welke kennis (o.a. van het vak in kwestie, maar bijvoorbeeld ook algemene didactische kennis, pedagogische content knowledge en evaluatiekennis) en welke strategische probleemaanpakken worden daarbij gehanteerd wanneer leerkrachten beslissingen nemen? Bij iets 'nieuws' als OGW zal men het dan tijdens de cognitieve taakanalyse wel eens dienen te worden over wat men onder 'goed OGW' verstaat.

Op basis van de resultaten van de cognitieve taakanalyse worden vervolgens op grond van een expliciete leerpsychologische rationale leertaken ontworpen die representatief zijn voor de beoogde complexe OGW-praktijk. Die leertaken worden geordend van simpele naar complexe taken. De ondersteuning die bij het leren uitvoeren van deze leertaken gegeven wordt is op een gegeven niveau van complexiteit eerst uitgebreid, maar neemt af naarmate men de taken beter beheerst. De lerende wordt bij het leren verwerven van de taken steeds voorzien van informatie die belangrijk is voor het uitvoeren van routinematige (en daarbinnen de wel/niet te automatiseren deeltaken) en non-routine deeltaken (de cognitieve strategieën en mentale modellen die belangrijk zijn voor het nemen van beslissingen in uiteenlopende situaties). Omdat een lerende uit de leertaken die taken kan kiezen, die gezien zijn/haar competenties het meest relevant voor hem/haar zijn zou deze aanpak mogelijk ook een benadering kunnen bieden voor het omgaan met de eerder besproken competentieverschillen tussen leerkrachten. De 4-componenten instructional design werkwijze heeft zich in andere contexten waarin sprake is van complex leren (bijvoorbeeld bij de opleiding van medisch specialisten, luchtverkeersleiders etc.) reeds veelvuldig bewezen. Het lijkt me zeer de moeite waard om deze aanpak in het kader van OGW en andere vormen van complex leren in/voor het onderwijs (de opleiding van leerkrachten bijvoorbeeld) toe te passen.

Paragraaf 2.1 over de factoren die een rol spelen bij goal setting en feedbackbenutting (en dus bij OGW) overziend, kan gesteld worden dat daarbij een veelheid aan factoren een rol speelt (zie onderstaande Figuur 9) en dat het antwoord op de vraag of OGW effectief is zal afhangen van de keuzes die gemaakt worden ten aanzien van die factoren. Wie krijgt bijvoorbeeld welk type feedback en hoe frequent? Wordt feedback gecombineerd met het stellen van doelen die SMART en uitdagend zijn? Zo ja, gebeurt dat alleen op het niveau van de klas, of ook op school- en bestuursniveau?





Figuur 9: Factoren die de benutting en effecten van feedback beïnvloeden

Worden de doelen zo opgesteld, dat betrokkenen zich daar echt geïmmiteerd aan voelen? Hoe competent zijn de feedbackontvangers om de gestelde doelen op basis van de feedback te realiseren? Ook zal de schoolcontext qua competenties, maar ook qua personeelsmobiliteit, beschikbare middelen e.d. een succesvolle invoering van OGW meer of minder waarschijnlijk maken. Het afstemmen van de inhoud van OGW-professionaliseringsinterventies op de, variërende, schoolcontext is daarom erg belangrijk. En, last but not least, zal een sterk leerpsychologisch fundament onder een OGW-professionaliseringsinterventie, waarin rekening wordt gehouden met wat het verwerven van elk van de OGW-vaardigheden, attitudes en -kennis veronderstelt, de kans op de effectieve invoering van OGW vergroten.

## 2.2 Wat weten we over de effectiviteit van OGW?

Van de kerningrediënten van OGW, feedback en goal setting is, zoals aangegeven, veelvuldig aangetoond dat ze prestatieverbeterend kunnen werken. Die ingrediënten worden bij OGW op een bepaalde manier ingevuld en sommige invullingen zijn hoogstwaarschijnlijk effectiever, dan andere. Onze kennis van effectieve OGW-aanpakken groeit weliswaar, maar is nog beperkt. De uitdaging ligt daarin, om, rekening houdend met de in paragraaf 2.1 besproken beïnvloedende factoren OGW-interventies te ontwerpen en in te voeren, en die grondig te evalueren op hun effecten. Daarbij kan worden voortgebouwd op de resultaten van reeds verricht OGW-onderzoek. Ik geef hier daarom een overzicht van de inhoud en resultaten van twee verschillende Nederlandse OGW-interventies: de Streef-interventies van de Rijksuniversiteit Groningen en de Focus-interventies die de afgelopen jaren door de Universiteit Twente zijn uitgevoerd.

De Rijksuniversiteit Groningen heeft in het Streef-project (120 scholen) samen met het Cito en twee onderwijsadviesorganisaties leerkrachten getraind en begeleid in een aantal OGW-trajecten (Doolaard, 2013a). In het eerste traject *Werken met referentieniveaus* stonden in vier bijeenkomsten de landelijk vastgestelde referentieniveaus en hun betekenis voor de eigen lespraktijk centraal. In het tweede traject *Gebruik maken van opbrengsten* zijn leerkrachten getraind in analysevaardigheden (leerlingvolgsystemenanalyses), het nemen van instructiebeslissingen op basis van de analyseresultaten, en de uitvoering van die beslissingen in hun eigen groep (negen bijeenkomsten).

Het derde traject *Streef middenbouw* besloeg de hele cyclus van opbrengstgericht werken: van analyse naar doelen stellen tot en met het uitvoeren en evalueren van het aangepaste onderwijs in de groep.

*Werken met referentieniveaus* liet zien dat leerkrachten slecht op de hoogte waren van de hoogte en de inhoud van de referentieniveaus. Vooral het hoogste niveau stelden zij desgevraagd veel hoger, dan wettelijk vastgesteld. Deelname aan dit deelproject leidde tot meer begrip van de (implicaties van de) referentieniveaus en ook tot meer realistische verwachtingen van leerlingen. Projectdeelname had geen effect op leerlingprestaties (Doolaard, 2013a).

In *Gebruik maken van opbrengsten* bleken leerkrachten hun analysevaardigheden (leerlingvolgsteemanalyses) behoorlijk te kunnen verbeteren en daar ook beter op te kunnen reflecteren. Hoewel de prestaties van leerlingen in de breedte een verschuiving in positieve richting lieten zien was het effect (net) niet statistisch significant en doofde het in het jaar na projectdeelname uit. Scholen gaven zelf aan een stok achter de deur nodig te hebben (Doolaard, 2013b).

*Streef middenbouw* liet op begrijpend lezen stevige effecten zien; leerlingen in groep 4 en 5 (420 leerlingen van 33 leerkrachten) presteerden 0.37 van een standaarddeviatie beter, dan de leerlingen in de m.b.v. propensity score matching gevormde controlegroep van 399 leerlingen (van Kuijk, 2014). Van differentiële effecten voor initieel hoog dan wel laag presterende leerlingen bleek geen sprake te zijn. Voor rekenen (Ritzema, 2015) werden geen effecten op leerprestaties gevonden (38 getrainde leerkrachten van groep 4 en 5; 527 leerlingen in de experimentele groep en eveneens 527 in de controlegroep). Getrainde leerkrachten hadden wel meer aandacht voor opbrengsten en waren ook meer doelgericht. Hogere doelen bleken bovendien samen te gaan met hogere leerprestaties (dit gold ook voor begrijpend lezen). Opmerkelijk was dat in klasobservaties de op instructie gerichte trainingscomponenten (directe instructie, modelleren, differentiatie) weinig effect op de instructie bleken te hebben gehad. Differentiëren gebeurde nog maar zeer beperkt en als leerkrachten de instructie in de klas wel differentieerden, dan deden ze dit voor zeer zwakke leerlingen. Goede leerlingen kregen geen extra instructie en de klassikale fase werd voor hen ook niet verkort.

Het Focus-project van de Universiteit Twente (zie <http://project-focus.gw.utwente.nl/>) betreft ook een interventie gericht op opbrengstgericht werken. In de periode van 2010 tot 2015 is deze in drie varianten aangeboden: de eerste keer als tweejarige variant, waarbij in beide jaren een gedeelte van het team werd getraind (Focus I, 43 scholen), daarna als tweejarige teamtraining (tweemaal: Focus II, 67 scholen; Focus III, 44 scholen) en tot slot als éénjarige training voor individuele leerkrachten (Focus IV, 71 leerkrachten).

Na de Focus I interventie werd geen verschil in leerprestaties gevonden tussen de experimentele groep van getrainde scholen en de met propensity scores verkregen groep van controlescholen, die de training niet hadden gevolgd. Van de teamtrainingen bleken de tweejarige interventies voor *volledige* teams over de hele interventieperiode de sterkste effecten op leerlingprestaties te hebben. In Focus II (ongeveer 7.500 leerlingen) werd een effect van gemiddeld één maand extra leerwinst voor het vak rekenen gerealiseerd (in 53 scholen, Focus II; van Geel, Keuning, Visscher & Fox, in druk). In Focus III konden scholen qua schoolvakken zelf hun professionaliseringstraject samenstellen: 35 scholen werkten aan OGW bij het vak rekenen (4.755 leerlingen) terwijl 28 scholen hun OGW-activiteiten op spelling richtten (3.975 leerlingen). Ook na Focus III werd zowel voor rekenen als voor spelling een gemiddeld effect van één maand extra leerwinst gedurende de hele interventieperiode vastgesteld

(Keuning, van Geel, Visscher & Fox, manuscript ter beoordeling ingediend). In beide projecten was het positieve effect op rekenprestaties het grootst voor laag-SES leerlingen (met een leerlinggewicht van 1.2) en voor hoog-SES leerlingen. *Scholen* met veel gewichtenleerlingen bleken het meeste profijt te hebben gehad van de Focus II- en de Focus III-interventie gericht op rekenen (vergeleken met scholen met een gemiddeld aantal, of weinig gewichtenleerlingen). Voor het vak spelling werden geen interventie-effecten gevonden, hier was sprake van een overall effect op leerlingen van gemiddeld één maand extra leerwinst gedurende de hele tweejarige interventieperiode.

Deelnemers aan de Focus II- en Focus III-interventie waren na afloop ook beter in staat om de prestatiegegevens van leerlingen, groepen en de hele school te interpreteren (1.182 respondenten, 101 scholen, Focus II en Focus III; Van Geel, Keuning, Visscher & Fox, manuscript ter beoordeling ingediend-a). Leerkrachten bleken gedurende de interventie doelgerichter te zijn gaan werken, en zij daagden hun leerlingen ook meer uit in hun lessen (37 scholen, Focus III, 209 leerkrachten van groep 4 t/m 8; Keuning, van Geel, Visscher & Fox, manuscript ter beoordeling ingediend). Verder bleek dat de samenwerking binnen teams intensiever was geworden, vooral wat betreft het bespreken van de prestaties van leerlingen (in 32 scholen, Focus III; Keuning, van Geel, Visscher & Fox, manuscript ter beoordeling ingediend) en dat de cultuur in de deelnemende scholen na afloop van de interventie meer te typeren was als een 'data culture' (Van Geel, Keuning, Visscher & Fox, manuscript ter beoordeling ingediend-b).

De intensieve begeleiding van 25 individuele leerkrachten gedurende een periode van 1 jaar in Focus IV (van der Scheer en Visscher, in druk), waarbij ook in de klas begeleiding en feedback aan leerkrachten werd gegeven, resulteerde erin dat leerkrachten:

- in de experimentele groep in vergelijking met leerkrachten in de controlegroep een extra leerwinst van twee maand voor rekenen realiseerden voor de verlengde-instructiegroep leerlingen (de zwakkere leerlingen).
- meer opbrengstgericht les gaven, dan vóór de training; ze bespraken en evalueerden namelijk vaker het lesdoel tijdens hun lessen en ze gaven ook vaker extra instructie aan de leerlingen in de verlengde instructiegroep;
- na de training significant meer vertrouwen hadden in hun eigen vaardigheid (self-efficacy) om leerlingen bij de les te betrekken en om aan hen passende instructiestrategieën aan te bieden (dit effect was een jaar na de training nog steeds aanwezig) (van der Scheer & Visscher, manuscript ter beoordeling ingediend).

De zes interventies en bijbehorende evaluatiestudies variëren qua:

- interventie-inhoud en interventieduur,
- onderzoekopzet (en daarmee ook wat betreft de basis voor causale claims),
- interventiedoelgroep: individuele leerkrachten, dan wel hele schoolteams (alle teamleden tegelijk, dan wel getrapt),
- vakken: vooral rekenen, maar ook begrijpend lezen en spelling,
- leerjaren: de groepen 4 en 5, groep 6, en groep 3 tot en met 8,
- schaal waarop aan OGW gewerkt is: de range van 420 leerlingen - 7500 leerlingen.

Tabel 1: De kenmerken en resultaten van de Streef & Focus OGW-interventies

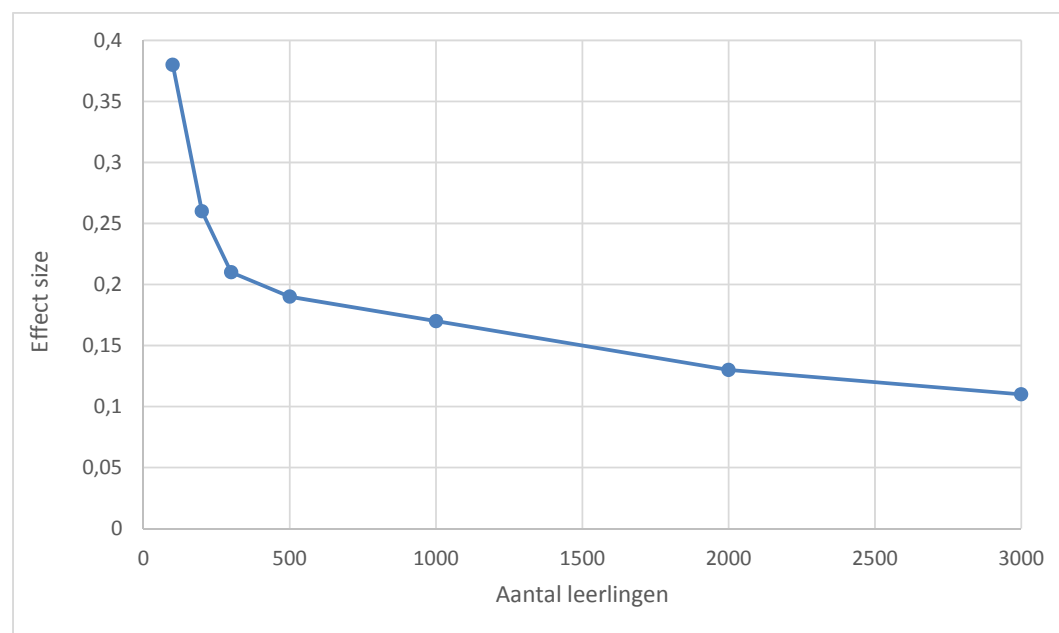
	Streef		Focus			
	I	II	I	II	III	IV
Gericht op	Klas	Klas	Klas	Klas en school	Klas en school	Klas
Onderzoeks-design	Propensity score matching	Propensity score matching	Propensity score matching	Multiple single subject	Multiple single subject	Random experiment
Duur	1 jaar	1 jaar	1 + 1 jaar	2 jaar	2 jaar	1 jaar
Deelnemers						
- Schoolleider	✓	✓	✓	✓	✓	
- Intern begeleider	✓	✓	✓	✓	✓	
- Leerkracht	✓	✓	✓	✓	✓	✓
Vakgebied	BL	RW	RW	RW	RW & SP	RW
Leerjaar	4 + 5	4 + 5	3 t/m 8	3 t/m 8	3 t/m 8	6
Aantal leerlingen	420	527	5913 5970	7500	RW: 4750 SP: 4000	673
Overall effect	0.37			1 mnd	RW: 1 mnd SP: 1 mnd	
Differentiële effecten				laag-SES & hoog-SES	RW: laag-SES & hoog-SES	2 mnd voor zwakker presterende leerlingen

BL = Begrijpend Lezen; RW = Rekenen-Wiskunde; SP = spelling

Als we ons beperken tot de effecten op leerprestaties, dan blijkt dat in vier van de zes OGW-interventies respectabele effecten op leerprestaties werden gerealiseerd. Die effecten zijn gemeten met een onafhankelijke, gestandaardiseerde test. Met dergelijke testen is het aanzienlijk moeilijker om effecten aan te tonen, dan met door onderzoekers zelf ontwikkelde toetsen (die doorgaans nauw aansluiten bij de interventie-inhoud, Cheung & Slavin, 2015). De positieve effecten op leerprestaties gelden zowel voor rekenen, begrijpend lezen als voor spelling. Dat bovendien voor de leerjaren 3 tot en met 8 effecten op leerprestaties zijn aangetoond is een belangrijke bevinding, omdat in de kernvakken in de lagere leerjaren meer wordt geleerd (daar is dus ook eerder winst te boeken met interventies), dan in hogere leerjaren. Hierbij moet wel worden aangetekend dat in de Focus II studie en de Focus III studie *gemiddelde* effecten over de leerjaren 3 tot en met 8 zijn berekend, nadere analyse moet nog uitwijzen hoe groot de effecten voor elk van die leerjaren zijn. In drie van de vier studies met positieve effecten op leerprestaties blijken vooral bepaalde groepen leerlingen van de interventies te profiteren: de zwakker presterende leerlingen, dan wel leerlingen met een hoge, of met een lage SES. OGW is bedoeld voor de maximalisering van de prestaties van *alle* leerlingen, dit blijkt dus nog niet altijd (wel in de Streef I interventie en de Focus III interventie gericht op spelling) het geval te zijn. Het zou mooi zijn als de grote groep *gemiddeld* presterende leerlingen ook meer profiteert van

OGW. Misschien richten leerkrachten zich tot nu toe vooral op leerlingen die qua prestaties opvallen en worden vooral voor hen interventies bedacht en blijven daardoor gemiddelde leerlingen de gemiddelde/gewone aanpak van de leraar krijgen.

Van belang is ook dat de effecten op leerprestaties werden vastgesteld in zowel kleinschalige (Streef I en II, Focus IV) als in grootschalige implementaties (Focus II en Focus III). Cheung & Slavin (2015) stelden op basis van de analyse van 645 interventiestudies een sterk verband vast tussen de schaal van implementatie en de omvang van de interventie-effecten. Figuur 10 laat zien dat grote effecten op leerprestaties (gemiddeld 0,38) werden gevonden wanneer in interventies met weinig leerlingen (+/- 100) werd gewerkt. Bij 200 leerlingen is het gemiddeld effect al tot 0,26 gedaald, bij 500 leerlingen daalt het effect verder tot 0,19, om vervolgens naar 0,11 te dalen in interventies waarin met 3000 leerlingen wordt gewerkt. In Focus II en Focus III werden voor aanzienlijk grootschaligere implementaties (4000, 4750 en 7500 leerlingen) effecten van gemiddeld een maand extra leerwinst voor alle betrokken leerlingen vastgesteld. Een maand leerwinst voor 4000 - 7500 leerlingen gedurende een periode van twee jaar is substantieel. Uiteindelijk gaat het om de effecten van grootschalige implementaties, want het OGW-beleid is bedoeld om landelijk te worden geïmplementeerd en de hoofdvraag is hoeveel in grootschalige implementaties overblijft van de effecten die in kleinschalige implementaties worden gerealiseerd. Bij kleinschalige interventies kan men interventies bijna perfect implementeren (men spreekt wel van 'super realization', Slavin & Smith, 2011). Bij opschaling heeft men alles niet meer volledig onder controle en krijgt men te maken met de imperfecties van werken 'in the real world'.



Figuur 10: Interventie-effecten in relatie tot implementatieschaal (Cheung & Slavin, 2015)

Naast de Streef- en Focus-studies zijn enkele andere Nederlandse studies ook interessant voor ons inzicht in de effectiviteit van OGW. Faber & Visscher (2014) voerden een systematische review uit van alle (inter)nationale studies waarin de effecten van digitale leerlingvolgsystemen (DLVS-en) op leerprestaties werden onderzocht. Wereldwijd bleken 15 onderzoeken te voldoen aan strenge

inclusiecriteria (o.a. alleen experimenteel onderzoek). De resultaten lieten zien dat het gebruik van een DLVS in het kader van onderwijs aan een klein aantal leerlingen een effect van 0,40 standaarddeviatie op leerprestaties had. Dit is een relatief groot effect in vergelijking met andere interventies die gericht zijn op *kleine aantallen leerlingen*. In studies waarin DLVS-en gebruikt werden voor het onderwijs aan alle leerlingen binnen een school, of schoolbestuur werd een effectgrootte van 0,06 gevonden. In vergelijking met andere interventies gericht op *scholen* is dit een iets lager, dan gemiddeld effect op leerprestaties. De volgende factoren bleken verder het effect van DLVS-gebruik op leerresultaten te bevorderen:

- een hoge *frequentie* waarmee het DLVS feedback verstrekt aan leerkrachten en/of scholen,
- DLVS-en die leerkrachten ook *advies* geven m.b.t. de instructie en de opdrachten voor leerlingen,
- een *interventie* met tenminste maandelijkse implementatie-activiteiten.

Veldhuis (2015) rondde recentelijk zijn dissertatieonderzoek af naar het gebruik door leerkrachten van toetstechnieken<sup>4</sup> waarvan de effectiviteit in internationaal onderzoek (Veldhuis, 2015) reeds vaak is aangetoond, maar die in de onderwijspraktijk nog weinig gebruikt worden. In een nascholingsproject werden 30 leerkrachten (616 leerlingen) begeleid in het gebruik van deze toetstechnieken. Leerkrachten waren enthousiast: de toetstechnieken bleken gemakkelijk te gebruiken te zijn en waardevolle informatie voor de instructie op te leveren. In het onderzoek was sprake van drie experimentele condities (één, twee, dan wel drie workshops om met de toetstechnieken te leren werken) en één controleconditie. De leerlingen van leerkrachten die drie workshops volgden gingen het meest vooruit op de leerlingvolgysysteemtoetsen rekenen/wiskunde ( $d = 0.55$  in vergelijking met  $d = 0.39$  voor de controleconditie).

De in deze paragraaf beschreven studies leveren het volgende beeld op:

- De analyse en interpretatie van leerlingvolgysysteemdata (blok 1 in Figuur 6) blijkt goed trainbaar te zijn (Staman, Visscher & Luyten, 2014) hetgeen een belangrijk resultaat is, omdat in de literatuur (zie o.a. Hamilton et al., 2009) gesteld wordt dat het gebruik van dit type data en de statistiek die eraan ten grondslag ligt een probleem vormen voor practici.
- Het lijkt erop dat de in vier OGW-interventies gevonden positieve effecten op leerprestaties vooral worden veroorzaakt, doordat leerkrachten en schoolteams leerden om de prestatiegegevens van hun leerlingen te analyseren, reflecteerden op de resultaten én op basis daarvan meer gefocust (al of niet aan de hand van expliciete doelen) en met inzet aan de leerlingprestaties werkten.
- De veranderingen in het instructiegedrag van leerkrachten in de klas lijken vooralsnog beperkt, zowel in de Streef- als in de Focus-projecten (vgl. ook Faber, Visscher en Schut, 2015). Hoewel er wel indicaties zijn van doelgerichtere en uitdagendere lessen vormen de vertaling van leerlingprestatiedata in 'remedies' (het beantwoorden van 'what-if-vragen' in het hoofd dan wel op papier) en de differentiatie van de instructie in de klas nog een probleem. We hebben in het onderzoek vooral geconstateerd dat er *kwantitatief* nog weinig sprake is van

---

<sup>4</sup> Bijvoorbeeld: de leerkracht vraagt: "Komt 44 in de tafel van 8 voor?". Idem: "Is het verschil tussen 15 en 7 groter dan 10?" Alle leerlingen antwoorden door een kaart met 'ja' of 'nee' omhoog te steken.

differentiatie. Het ontbreekt eigenlijk nog aan een instrumentarium waarmee de *kwaliteit* van de differentiatie kan worden vastgesteld: hoe goed sluiten de handelingen van de leerkracht aan bij wat een leerling gezien zijn prestatiebeeld nodig heeft?

- De frequentie en inhoud van de feedback aan leerkrachten (liefst inclusief advies over hoe de feedbackontvanger kan verbeteren) evenals de begeleiding van en training in feedbackgebruik blijken belangrijke moderatoren te zijn voor de effecten van prestatiefeedback op leerprestaties.
- Andere toetsen en toetstechnieken, dan leerlingvolgsystemen kunnen ook bijdragen aan meer op onderwijsbehoeften afgestemde instructie en daarmee aan hogere leerlingprestaties. De toetstechnieken, zoals onderzocht door Veldhuis (2015) zijn daarvan een mooi voorbeeld. Digitale toetsystemen á la Rekentuin, Snappet en Taalzee worden ook steeds meer gebruikt en zullen waarschijnlijk belangrijk worden in het onderwijs. Deze systemen bieden in vergelijking met de leerlingvolgsysteemfeedback meer 'rapid feedback'<sup>5</sup> (Yeh, 2008) aan leerlingen die opgaven maken én aan leerkrachten, die zo kunnen zien welke opgaven door welke leerlingen wel/niet beheerst worden. De systemen zijn vaak adaptief in die zin dat ze de leerlingopgaven differentiëren op basis van het vastgestelde niveau van een leerling. In die zin wordt door de digitale systemen in de opgaven naar leerlingen toe dus al meer gedifferentieerd. Onze lesobservaties maakten duidelijk dat leerkrachten bij Snappet nog weinig deden met de feedback die ze ontvingen (Faber & Visscher, 2016). Ook hier maakt het nog onvoldoende beheersen van de voor instructiedifferentiatie benodigde competenties een volledige feedbackbenutting waarschijnlijk nog niet mogelijk.

Nu de evidentie met betrekking tot de effectiviteit van OGW is besproken is het tijd om stil te staan bij het onderzoek dat ons op dit terrein verder kan brengen.

### **3. Toekomstig onderzoek naar OGW**

Van de in paragraaf 2.1 beschreven factoren die de benutting van feedback beïnvloeden (zie ook Figuur 9) dienen er in toekomstig OGW-onderzoek naar mijn mening twee met name aandacht te krijgen:

1. de professionalisering voor OGW en daarbij vooral voor het differentiëren op basis van prestatiefeedback;
2. het gebruik van 'rapid feedback' door leerkrachten, die onder andere met behulp van digitale toets- en feedbacksystemen in de klas verkregen kan worden.

Bij het eerste punt bestaat de uitdaging daarin nader te onderzoeken hoe leerkrachten, schoolteams en bestuurders goed toegerust kunnen worden voor het benutten van leerlingvoortgangsfeedback voor de verbetering van onderwijskwaliteit en leerprestaties. Willen niet alleen leerkrachten, maar ook scholen en hun besturen meer gaan opereren als lerende organisaties, die op basis van prestatiefeedback op hun eigen functioneren en presteren reflecteren en die systematisch zoeken naar manieren voor prestatieverbetering, dan zal dit voor veel scholen een cultuuromslag betekenen die om nieuwe competenties en dus om professionalisering vraagt. Tot nu toe zijn besturen nog

---

<sup>5</sup> Leerkrachten gebruiken ook methodegebonden toetsen, geven beurten en kijken gemaakte opgaven na.

nauwelijks systematisch getraind voor dit doel. Voor professionalisering op schoolniveau kennen we al wel voorbeelden (Focus II en Focus III), maar de meest trajecten zijn tot nu toe nog op leerkrachten gericht geweest. Daar is qua effecten op leerprestaties ook de grootste leerwinst te behalen, maar daar ligt ook de grootste uitdaging, omdat het tot nu toe nog niet goed lukt om leerkrachten toe te rusten voor differentiëren op basis van prestatievoortgangsinformatie.

In het voorgaande gaf ik reeds aan dat een probleem bij veel onderwijsbeleidsinitiatieven is dat ze globaal omschreven zijn (bijvoorbeeld: 'gebruik leerlingvolgsysteemdata voor het verzorgen van onderwijs op maat'). Die globale omschrijvingen zijn op tal van, uiteenlopende, manieren concreet in te vullen en ".....*many interventions thus leave much to be filled in by teachers and schools* (Cohen en Ball, 2001)". Omdat het beoogde nieuwe gedrag niet helder omschreven is kan er ook niet systematisch voor getraind worden. We hebben dus heldere definities van de gewenste competenties nodig.

Ik verwacht hierbij veel van de toepassing van het 4-componenten instructional design (4C/ID) model van Van Merriënboer en Kirschner (2013) in drie nieuwe onderzoeksprojecten waarin een grondige cognitieve taakanalyse de basis vormt voor leerkrachtprofessionaliseringstrajecten met een sterke leerpsychologische basis (zie ook 2.1.5 voor meer over 4C/ID).

Het is belangrijk dat we ons realiseren dat een systematische training van de vereiste complexe OGW-competenties essentieel is, maar dat het afleren van oude werkwijzen en het gedisciplineerd toepassen van de nieuwe werkwijzen ook van groot belang is. We weten op andere terreinen (een gezonde levensstijl bijvoorbeeld) best hoe 'het moet', maar passen die kennis niet altijd toe, omdat de minder gezonde levensstijl ook zo zijn aantrekkingskracht heeft, veel discipline vereist e.d. OGW veronderstelt een weloverwogen, nauwgezette werkwijze die niet alleen veel competenties maar ook de nodige energie en inzet vereist.

Ook is het goed dat we ons realiseren dat er grenzen aan professionalisering zijn. In die zin, dat het *totale* pakket aan kennis en vaardigheden dat OGW veronderstelt niet in 5-10 scholingsbijeenkomsten per schooljaar kan worden aangeboden: de vakinhoudelijke kennis van elk van de kernvakken, de algemeen-didactische kennis en vaardigheden en de benodigde evaluatie- en toetskennis. Daarom is het ook zo belangrijk dat de pre-service opleiding van leerkrachten (de pabo) daar grotendeels in voorziet, en dat de in-service professionalisering van leerkrachten daar op voort kan bouwen.

Het tweede, aan het begin van deze paragraaf genoemde punt dat voor toekomstig onderzoek belangrijk is betreft het gebruik van digitale of mondelinge (Veldhuis, 2015) 'rapid feedback' door leerkrachten. In toekomstig onderzoek worden idealiter de effecten onderzocht van rapid feedback in combinatie met de halfjaarlijkse feedback uit leerlingvolgsystemen én de training van leerkrachten om de rapid en slow feedback adequaat te benutten. Het halfjaarlijkse ijkpunt van het leerlingvolgsysteem blijft belangrijk, omdat dit aangeeft wat een leerling na een half jaar nog weet van de diverse tijdens dat half jaar behandelde onderwerpen. Deze feedback kent echter zijn beperkingen voor het *dagelijkse* opbrengstgericht handelen in de klas.

Toekomstig OGW-onderzoek is idealiter gebaseerd op een theory of action/improvement, waarbij Figuur 8 het uitgangspunt zou kunnen vormen. Eventueel kan tussen het tweede en derde blok nog



een extra blok 'geplande instructie' worden toegevoegd, zodat ook de interventie-effecten op lesvoorbereidingsactiviteiten (lesplan, groepsplan) bestudeerd kunnen worden.

Door niet alleen de effecten van een OGW-professionaliseringsvariant op leerprestaties te onderzoeken, maar alle schakels in de keten van Figuur 8 wordt inzicht verkregen in de kwaliteit waarmee elk van de schakels in de Figuur wordt uitgevoerd en hoe dat doorwerkt in de daarop volgende schakels. Dat kan verklaringen bieden voor het wel of niet optreden van de beoogde effecten op leerprestaties en tevens input vormen voor de ontwikkeling en implementatie van nieuwe OGW-professionaliseringsinterventies.

### *Nieuwe onderzoeksprojecten*

Een aantal recent gestarte, op differentiatie gerichte onderzoeksprojecten is hier noemenswaardig.

In een door het GION van de Rijksuniversiteit Groningen uitgevoerd NRO-project wordt de Nederlandse versie van het Amerikaanse Success for All programma (e.g. Slavin et al., 2009) geëvalueerd op leerprestatie-effecten. In Success for All is het frequent monitoren en bespreken van leerlingprestatiegegevens de basis voor het samenstellen van homogene, leerjaaroverstijgende leesgroepen en voor het formuleren van prestatiedoelen. In de homogene leesgroepen krijgen leerlingen elke dag 90 minuten lang leesles op hun eigen niveau. De rest van de dag werken leerlingen binnen hun leerjaar in heterogeen samengestelde groepen (qua prestaties, ras en geslacht) op basis van coöperatief leren. Inzicht in de effecten van de Success for All aanpak in de Nederlandse onderwijscontext is van belang en kan van betekenis zijn voor de vormgeving van nieuwe OGW-varianten.

De Lerarenopleiding van de Rijksuniversiteit Groningen onderzoekt in een andere, internationaal vergelijkende NRO-studie hoe in andere landen met het differentiatievraagstuk wordt omgegaan hetgeen een inspiratiebron kan vormen voor Nederlandse differentiatie-aanpakken.

De Universiteit Maastricht (Department of Educational Development & Research) en de Universiteit Twente (vakgroepen ELAN en OMD) werken samen in drie op differentiatie gerichte projecten. In het Match-project (NRO-studie) wordt op basis van het in het voorgaande kort getypeerde 4-componenten instructional design model van Van Merrienboer en Kirschner (2013) een diepgaande analyse gemaakt van differentiëren als complexe didactische vaardigheid in het primair onderwijs. De analyse wordt zowel uitgevoerd voor traditionele scholen voor primair onderwijs, als voor basisscholen die ICT intensief voor hun onderwijs gebruiken. In laatstgenoemde scholen wordt ook met per vak wisselende, leeftijdsoverstijgende niveaugroepen gewerkt en vormt het monitoren van de voortgang een uitdaging. Met behulp van het 4C/ID-model zal op basis van de resultaten van de cognitieve taakanalyse onder leerkrachten die differentiëren als didactische vaardigheid goed beheersen, in samenwerking met inspecteurs, pabo's, trainers en andere experts (zowel qua differentiatie als qua vakdidactiek) een differentiatietraining worden ontworpen en geïmplementeerd. De trainingseffecten op de differentiatievaardigheden van leerkrachten en op de prestaties van leerlingen zullen worden onderzocht.

In het tweede project wordt in samenwerking met de Onderwijsinspectie onderzocht hoe basisschoolleerkrachten getraind kunnen worden voor de benutting van digitale feedbacksystemen

die aan leerkrachten niet alleen 'slow feedback' (uit een leerlingvolgsysteem), maar ook 'rapid feedback' geven over de vorderingen van leerlingen. Ook hier speelt de differentiatie van de instructie een grote rol en worden de effecten van instructiedifferentiatie op leerprestaties onderzocht.

Dit geldt eveneens voor het derde project, waarin in samenwerking met het Ministerie voor OCW in het voortgezet onderwijs wordt onderzocht hoe leerkrachten getraind kunnen worden voor differentiatie op basis van de feedback die ze ontvangen op basis van door Havo-4 leerlingen gemaakte schoolexamens in combinatie met de Proeftuin van de SLO (een online omgeving waarin leerlingen examens kunnen oefenen, die zowel aan leerkrachten als leerlingen feedback over vorderingen geeft).

#### **4. Enkele aanbevelingen voor de bevordering van Opbrengstgericht werken in het primair onderwijs**

De Streef- en Focus-projecten laten zien dat OGW substantiele effecten op leerprestaties kan hebben en dat het zinvol is om de opbrengstgerichtheid binnen het onderwijs te vergroten. De volgende aanbevelingen kunnen naar mijn mening bijdragen aan de verdere invoering van OGW.

##### *Aanbeveling 1: De OCW-doelstelling realiseren*

In 2010 stelde het Ministerie voor OCW dat in 2018 negentig procent van de basisscholen aan alle vijf OGW-indicatoren zou moeten voldoen (Visscher & Ehren, 2011). Volgens de peilingen van de Onderwijsinspectie voldoet echter al jaren slechts 25% - 35% van de basisscholen aan deze vijf indicatoren. Het gaat om de volgende indicatoren:

1. scholen gebruiken een leerlingvolgsysteem voor het volgen van de prestaties en de ontwikkeling van leerlingen,
2. de leraren volgen en analyseren systematisch de voortgang in de ontwikkeling van de leerlingen,
3. de school evalueert regelmatig de effecten van de zorg,
4. de school evalueert jaarlijks de resultaten van leerlingen,
5. de school evalueert regelmatig het onderwijsleerproces (1 x per 4 jaar).

Gemiddeld voldoen scholen nu aan 3,8 indicator, maar het blijkt maar niet te lukken om dit naar vijf indicatoren te brengen. De indicatoren die door de overheid gebruikt worden zijn meer op schoolniveau geformuleerd, dan de wijze waarop OGW in deze oratie beschreven wordt. Het gaat bij de overheidsindicatoren in feite om basiscriteria voor een goede bedrijfsvoering binnen scholen. Met de overheid vind ik dan ook dat basisscholen hier zoveel mogelijk aan moeten voldoen. De overheid doet er goed aan aanvullende maatregelen te ontwikkelen om haar eigen doelstelling te realiseren.

### *Aanbeveling 2: Alleen evidence-based professionalisering voor OGW*

In het basisonderwijs wordt jaarlijks voor vele miljoenen geprofessionaliseerd. In 2014 werkte 78.000 fulltime equivalent aan leerkrachten in het primair onderwijs. Elke leerkracht en elk directielid kan jaarlijks op basis van de cao een bedrag aan professionalisering besteden. Op jaarbasis gaat het hier in totaal om een bedrag van 45 miljoen euro aan 'verplichte' professionalisering. Daarnaast besteden scholen op eigen initiatief jaarlijks nog vele miljoenen extra aan professionalisering.

Professionaliseringstrajecten voor scholen behoren naar mijn mening zoveel mogelijk aangetoond positieve effecten op leerprestaties te hebben. Dit geldt zeker ook voor OGW-trajecten. Waarom zouden scholen investeren in OGW-trajecten waarvan niet is aangetoond dat ze tot betere leerprestaties leiden, terwijl van andere OGW-interventies de positieve effecten op leerprestaties zijn aangetoond? Belangrijk is dat scholen de aangeboden professionaliseringstrajecten op dit punt kritisch beoordelen voordat ze beslissen over deelname aan een traject.

### *Aanbeveling 3: Moderniseer het pabo-onderwijs*

Pabo's moeten hard aan de slag met het moderniseren en optimaliseren van hun onderwijs voor aanstaande leerkrachten: het gebruik van digitale feedbacksystemen (leerlingvolgsystemen, maar zeker ook de systemen die in de klas 'rapid feedback' kunnen leveren en die al veel gebruikt worden in de onderwijspraktijk) qua gebruik in *enge* zin: hoe werken die systemen, wat kan ik eruit halen, en wat betekent de output precies?. Aanstaande leerkrachten behoren echter vooral ook te leren hoe ze deze systemen in *brede* zin kunnen gebruiken: het verwerven van de competenties voor het verzorgen van onderwijs op maat op basis van leerlingprestatiefeedback. Daar is veel te winnen, met name qua instructiedifferentiatie. Naarmate we daar beter in slagen zal de in-service scholing van leerkrachten minder nodig zijn en kunnen meer leerlingen sneller van onderwijs op maat profiteren. De academische pabo's zijn wat dit betreft al een stap in de goede richting, maar er is meer nodig, vooral de modernisering van de pabo-curricula op genoemde punten.

## Dankwoord

Graag wil ik het College van Bestuur van de Rijksuniversiteit Groningen bedanken voor het instellen van de bijzondere leerstoel 'Opbrengstgericht Werken in het Onderwijs' via de Stichting *Success for All Nederland* én voor het door hen in mij gestelde vertrouwen.

Roel Bosker is degene die het initiatief tot de instelling van de leerstoel heeft genomen en dit plan daarna tot uitvoering heeft gebracht, door alle stappen die daarbij horen met succes uit te voeren. Dat waardeer ik zeer. Het eerste jaar zit er inmiddels op en is naar mijn mening heel plezierig verlopen. Ik voel me steeds welkom en vind het prettig om in de Groningse context te werken aan onderwijs en onderzoek. Geleidelijk aan raak ik bij steeds meer activiteiten betrokken en ik ga ervan uit dat we de komende jaren een mooie, praktijkrelevante én wetenschappelijk interessante verdere invulling aan de leerstoel gaan geven.

Anja Knuver wil ik bedanken voor haar zeer waardevolle feedback op de conceptversies van zowel de schriftelijke oratie als van de presentatie voor de 8<sup>e</sup> december. Van feedback kan iets beter worden en dat geldt zeker voor jouw feedback!

Ik heb het zeer getroffen met mijn AIO's, zowel wat betreft de prettige samenwerking als wat betreft hun professionele inbreng. Die inbreng is heel belangrijk geweest bij de vormgeving van de Focus-interventies en bij het onderzoek naar de effecten ervan. Mede daardoor heb ik deze oratie kunnen schrijven. Daarvoor ben ik Laura Staman, Marieke van Geel, Trynke Keuning, Emmelien van der Scheer en Marjan Faber dan ook enorm erkentelijk. Trynke, Marieke en Emmelien wil ik bovendien bedanken voor hun hulp bij de voorbereiding van de presentatie van de oratie. Fijn dat jullie mijn ruwe versie van de presentatie omtoverden in een presentatie met een mooie lay-out, die ook voor niet-kleurenblinden prettig leest!

Deze oratie draag ik op aan mijn ouders. Ze vonden onderwijs en de opleiding van hun kinderen erg belangrijk en deden daarvoor alles wat binnen hun vermogen lag. Ze hadden mijn benoeming op deze leerstoel op het terrein van opbrengstgericht werken in het onderwijs prachtig gevonden.

Ik heb gezegd.

## Literatuur

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Blok, H., Ledoux, G., & Roeleveld, J. (2015). Opbrengstgericht werken in het primair onderwijs: een effectieve weg naar onderwijsverbetering? *Pedagogische Studiën*, 92, 167-178.
- Cheung, A., & Slavin, R.E. (2015). *How Methodological Features Affect Effect Sizes in Education*. Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education.
- Cohen, D.K., & Ball, D.B., (2001). Making change: instruction and its improvement. *Phi Delta Kappan*, 83 (1), 73-77.
- Desimone, L. M. (2009). Improving impact studies of teacher's professional development: toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199. doi: 10.3102/0013189X08331140
- Deunk, M. I., Maslowski, R., van Kuijk, M., & Doolaard, S. (2015). *Scholen in verandering*. Groningen: GION.
- Doolaard, S. (2013a). *Het streven naar kwaliteit in scholen voor primair onderwijs*. Groningen: GION.
- Doolaard, S. (2013b). *Effecten van het trainings- en begeleidingstraject Streef: gebruik maken van opbrengsten*. Groningen: GION.
- Faber, J. M., & Visscher, A. J. (2014). *Digitale leerlingvolgsystemen: een review van de effecten op de leerprestaties*: Universiteit Twente (in opdracht van Kennisnet). Geraadpleegd via URL <http://www.kennisnet.nl/onderzoek/nieuws/digitale-leerlingvolgsystemen-een-review-van-de-effecten-op-leerprestaties>
- Faber, J. M., & Visscher, A. J. (2016). *De effecten van Snappet. Effecten van een adaptief onderwijsplatform op leerresultaten en motivatie van leerlingen*. In opdracht van Kennisnet. Enschede: Universiteit Twente, vakgroep ELAN.
- Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/wwc/publications/practiceguides>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. doi: 10.3102/003465430298487
- Ikemoto, G. S., & Marsh, J. A. (2007). Cutting Through the 'Data-Driven' Mantra: Different Conceptions of Data-Driven Decision Making. In P. A. Moss (Ed.), *Evidence and Decision Making* (Vol. 106, pp. 105-131). Malden, MA: Blackwell Publishing Inc.
- Inspectie van het Onderwijs (2010). *Opbrengstgericht werken in het basisonderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2014). *De Staat van het Onderwijs. Onderwijsverslag 2012/2013*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2015). *De Staat van het Onderwijs. Onderwijsverslag 2013/2014*. Utrecht: Inspectie van het Onderwijs.
- Keuning, T., Van Geel, M., Visscher, A., Fox, J. P., & Moolenaar, N. M. (in druk). The Transformation of Schools' Social Networks During a Data-Based Decision Making Reform. *Teachers College Record*.
- Keuning, T., Van Geel, M., Visscher, A. J., & Fox, J. P. (2015). *The Development of Teaching Quality During a DBDM-Intervention: A Pre-post Multilevel IRT Approach*. Manuscript ingediend voor beoordeling.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation. *American Psychologist*, 57(9), 705-717.
- Maslowski, R., Deunk, M. I., van Kuijk, M., & Bijlsma, H. (2015). *Adaptief onderwijsbeleid: Een verkenning onder scholen en besturen naar de mogelijkheden voor het voeren van een adaptief onderwijsbeleid*. Groningen: GION.
- Ministerie voor Onderwijs, Cultuur & Wetenschap (2007). *'Scholen voor Morgen'*. Den Haag: Ministerie voor Onderwijs, Cultuur & Wetenschap.
- Ministerie voor Onderwijs, Cultuur & Wetenschap. (2010). *Actieplan Basis voor presteren*. Den Haag: Ministerie voor Onderwijs, Cultuur & Wetenschap.
- Ritzema, L. (2015). *Professional development in data use: the effects of primary school teacher training on teaching practices and students' mathematical proficiency*. Groningen: GION.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153-189.
- Slavin, R.E. (2015). Response to Intervention and Bob's Law. *HuffPost Education*, November 19.

- Slavin, R.E. & D. Smith (2011). The relationship between sample sizes and effect sizes in systematic reviews in education. In W.P.Vogt (Ed.), *Sage Quantitative Research Methods* (Volume III: Alternatives to Hypothesis Testing, pp. 169 – 180). London: Sage Publishers Ltd.
- Slavin, R. E., Cheung, A., Holmes, G., Madden, N. A., & Chamberlain, A. (2013). Effects of a Data-Driven District Reform Model on State Assessment Outcomes. *American Educational Research Journal*, 50(2), 371-396. doi: 10.3102/0002831212466909
- Slavin, R.E., Madden, N.A., Chambers, B., & Haxby, B. (2009). *Two million children: Success for All*. Thousand Oaks, CA: Corwin.
- Staman, L., Visscher, A. J., & Luyten, H. (2014). The effects of professional development on the attitudes, knowledge and skills for data-driven decision making. *Studies in Educational Evaluation*, 42, 79-90. doi: 10.1016/j.stueduc.2013.11.002
- Timperley, H. (2008). *Teacher professional learning and development*. Brussels: International Academy of Education/International Bureau of Education.
- van der Scheer, E. A., & Visscher, A. J. (in druk). *Effects of a Data-Based Decision Making Intervention for Teachers on Students' Mathematical Achievement*.
- van der Scheer, E. A., & Visscher, A. J. (2015). *Effects of an Intensive Data-Based Decision Making Intervention on Teacher Efficacy Teaching and Teacher Education*. Manuscript ingediend ter beoordeling.
- van Geel, M., Keuning, T., Visscher, A. J. & Fox, J. P. (in druk). Assessing the Effects of a School Wide Data-Based Decision Making Intervention on Student Achievement Growth in Primary Schools. *American Educational Research Journal*.
- van Geel, M., Teunis, B., & Visscher, A. J. (2015). *School characteristics that influence the implementation of a data-based decision making intervention*. Manuscript ingediend ter beoordeling.
- van Geel, M., Keuning, T., Visscher, A. J. & Fox, J. P. (2015a). Changes in Educators' Data Literacy during a Data-Based Decision Making Intervention. Manuscript ingediend ter beoordeling.
- van Geel, M., Keuning, T., Visscher, A. J. & Fox, J. P. (2015b). Development of Schools' Data Culture during a School-Wide Data-Based Decision Making Intervention. Manuscript ingediend ter beoordeling.
- van Kuijk, M. (2014). *Raising the bar for reading comprehension*. Groningen: GION.
- van Merriënboer, J.J.G. & Kirschner, P.A. (2013). *Ten steps to complex learning*. New York: Routledge.
- van Veen, K., Zwart, R., & Meirink, J. (2011). What Makes Teacher Professional Development Effective? In M. Kooy & K. Van Veen (Eds.), *Teacher learning that matters: International perspectives* (pp. 3-21). London: Routledge.
- Veldhuis, M. (2015). *Improving classroom assessment in primary education*. Utrecht: Freudenthal Institute for Science and Mathematics Education.
- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, 24, 80-91.
- Visscher, A. J., & Coe, R. (2002). *School Improvement through Performance Feedback*. Lisse: Swets & Zeilinger.
- Visscher, A., Ehren, M. (2011). *De eenvoud en complexiteit van opbrengstgericht werken*. Rapport voor de Kenniskamer van het Ministerie voor Onderwijs Cultuur & Wetenschap. Den Haag: Ministerie voor Onderwijs Cultuur & Wetenschap.  
<https://www.rijksoverheid.nl/documenten/rapporten/2011/07/13/de-eenvoud-en-complexiteit-van-opbrengstgericht-werken>
- William, D. (2006). *Assessment for Learning: why, what and how*. Edited transcript of a talk given at the Cambridge Assessment Network Conference on 15 September at the Faculty of Education, University of Cambridge.
- Yeh, S. S. (2008). The cost-effectiveness of comprehensive school reform and rapid assessment. *Education Policy Analysis Archives*, 16(13).
- Yoon, K. S., Duncan, T., Lee, S. W.Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Geraadpleegd via URL <http://ies.ed.gov/ncee/edlabs>.